DOCUMENT RESUME

ED 379 806 EA 026 564

AUTHOR Sperry, David J.; And Others
TITLE Educator Evaluation and the Law: Confusion in Purpose. Policy Studies in Education.
INSTITUTION Utah Univ., Salt Lake City. Utah Education Policy Center.
PUB DATE 1 Oct 92
NOTE 32p.
AVAILABLE FROM Utah Education Policy Center, School of Education, University of Utah, Salt Lake City, UT 84112 ($5; contact publisher for quantity discounts).
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Legislation; Elementary Secondary Education; *Evaluation Problems; *Evaluation Utilization; Faculty Evaluation; Instructional Effectiveness; Reliability; *State Legislation; Teacher Behavior; *Teacher Evaluation; Validity
IDENTIFIERS *Utah

ABSTRACT

Many states have a disconnected body of laws regarding educator evaluation. Uncoordinated legislation can create problems with respect to evaluation, appropriate standards and methods to be employed, and implementation strategies. This paper describes Utah's educator-evaluation legislation to illustrate this phenomenon. Four statutory provisions are examined--the Utah Orderly Termination Procedures Act, the Educational Professional Practices Act, the Career Ladder Act, and the Educator Evaluation Act. The analysis focuses on the congruency between the intent of state legislative provisions and evaluation purposes. The paper also examines the relation of technical soundness to evaluation purposes. Effective evaluation programs include several elements that are difficult to implement when legislation is assembled in a piecemeal manner: (1) a crisp and functional articulation of the evaluation purposes; (2) evaluation methods that match assessment purposes; (3) technical soundness, such as reliability and validity; and (4) feasibility and practicability. Utah's major educator evaluation laws illustrate a number of difficulties commonly found nationwide--vague legislative language, variable terminology, unclear assessment standards, intersecting and incompatible laws, mixed purposes, and unclear use of the term "competence." Recommendations for administrators and policymakers are included. (LMI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 379 806

POLICY STUDIES IN EDUCATION

Educator Evaluation and the Law: Confusion in Purpose

David J. Sperry, Diana G. Pounder and Clifford J. Drew

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

[X] This document has been reproduced as received from the person or organization originating it.

[ ] Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Sperry

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Published by the

UTAH EDUCATION POLICY CENTER
The Graduate School of Education
University of Utah
Salt Lake City, Utah

BEST COPY AVAILABLE

EA 026 564

EDUCATOR EVALUATION AND THE LAW:

CONFUSION IN PURPOSE

David J. Sperry, Diana G. Pounder,

and

Clifford J. Drew

Utah Education Policy Center

Graduate School of Education

Milton Bennion Hall

The University of Utah

Salt Lake City, UT 84112

October 1, 1992

AUTHORS

Dr. Sperry is Chair and Associate Professor, Department of Educational Administration, University of Utah.

Dr. Pounder is Associate Chair and Associate Professor, Department of Educational Administration, University of Utah.

Dr. Drew is Associate Dean for Research in the Graduate School of Education and Professor of Special Education and Educational Psychology, University of Utah.

EDUCATOR EVALUATION AND THE LAW: CONFUSION IN PURPOSE

HIGHLIGHTS

STATUTORY DEVELOPMENT PERTAINING TO EDUCATOR EVALUATION IN UTAH PRIMARILY INVOLVES FOUR LAWS ENACTED OVER A 14-YEAR PERIOD

* Because these laws were assembled over time and in a piecemeal fashion, a disconnected body of laws and amendments has resulted which is creating problems for educator evaluation in schools.
* These laws appear to have overlapping purposes, inconsistent purposes, and ambiguities and inconsistency in terminology, presenting school districts with a number of implementation dilemmas.
* Appropriate standards, methods of evaluation, and implementation strategies are often unclear or inconsistent with the implied purpose for evaluation.

EFFECTIVE EVALUATION PROGRAMS INCLUDE SEVERAL ELEMENTS THAT ARE DIFFICULT TO IMPLEMENT WHEN LEGISLATION IS ASSEMBLED IN A PIECEMEAL MANNER

* Effective evaluation requires a crisp and functional articulation of the purposes for evaluation.
* Effective evaluation must match evaluation types and methods to the purposes for which the assessment is being undertaken--such as staff development, promotion, or termination.
* Technical soundness of evaluation cannot be adequately addressed without the evaluation purpose.
* Effective evaluation must be practicable within the parameters of available resources.

UTAH'S MAJOR EDUCATOR EVALUATION LAWS ILLUSTRATE A NUMBER OF DIFFICULTIES COMMONLY FOUND ACROSS THE COUNTRY

* Legislative language often addresses purpose only vaguely if at all.
* Terminology varies among related laws.
* Assessment standards and procedures are unclear and subject to multiple interpretations.
* Purposes, where addressed, appear to intersect and even be incompatible between laws.
* Different purposes are mixed even within the same legislation, resulting in certain circumstances where incompatible evaluation procedures may result from the same law.
* Combinations of evaluation of instructional performance and behavioral misconduct matters occur in some legislation with reference to the term "competence" for both areas.

RECOMMENDATIONS FOR LOCAL ADMINISTRATORS ACTING UNDER EXISTING LAWS

* Keep purposes of each and every evaluation clear. (i.e., is it formative, summative, or both; is it to promote professional growth, to determine eligibility for career ladder monies, to be used for possible termination action, or some combination?)
* Match an evaluation or an evaluation system to the intended purpose(s).
* Make a "good faith effort" to identify and comply with all procedural and substantive requirements noted in the Utah Orderly Termination Procedures Act, the Career Ladder Act and the Educator Evaluation Act.
* Because current legislative language requires "a reliable and valid evaluation consistent with generally accepted professional standards for personnel evaluation systems" and given the fact that no single set of standards enjoys widespread acceptance, local school boards should adopt a set of "defensible" standards which enjoys the support of both local teachers and administrators.

* Establish and maintain a continuous training and monitoring system of all evaluation instruments and procedures. Data need to be gathered and evaluated constantly to preserve the integrity of the system. This will require the use of highly trained specialized personnel.

RECOMMENDATIONS FOR STATE POLICY-MAKERS INTERESTED IN CHANGING EXISTING LAWS

* Incorporate critical elements and primary purposes of the four evaluation-related acts into an omnibus educator evaluation act to reduce the ambiguity, overlap, and potential contradictions currently created by terminology and provisions of the current four separate acts. For instance, the omnibus act could include:
  a) a provision for formative evaluation and staff development (similar to some aspects of the existing Educator Evaluation Act);
  b) a provision for summative evaluation or employment retention/termination decisions including due process provisions (similar to some aspects of the Educator Evaluation Act and primary aspects of the Orderly Termination Act);
  c) a provision for rewarding performance (similar to primary aspects of the Career Ladder Act);
  d) a provision for revocation of licensure (a primary aspect of the Educational Professional Practices Act).

  Most importantly, this omnibus act should have clear and consistent use of terms and make provision for the full array of personnel functions and decisions that represent the various purposes for educator evaluation.

* The degree of specificity in a newly created omnibus act should be somewhat less than currently dictated by similar acts. In particular, the specific methods, time-frames, etc. that are appropriate for each major type of evaluation should be determined by a joint committee of teachers and administrators at the local district level. The omnibus legislation should provide the basic purposes and broad legal

principles that must be implemented. The local districts should be granted the discretion to adopt implementation policies and practices that fit local district needs and professional norms, yet adhere to broad principles of fairness, accuracy, effectiveness, and practicability.

Introduction

Statutorily many states have gone through an evolutionary process regarding educator evaluation. This often results in a disconnected body of laws and amendments that create problems with respect to purposes of evaluation, appropriate standards and methods to be employed, and implementation strategies. The Utah case study presented in the current paper illustrates this phenomenon. As educator evaluation matures, statutorial efforts to look at educator evaluation in a comprehensive manner becomes critical. Effective evaluation programs include a number of elements that are difficult to implement when laws and regulations are enacted in a piecemeal manner. The first of these is a crisp, functional articulation of the purposes for evaluation and the corresponding types of evaluation. Second, evaluation must address the technical soundness of assessment such as reliability and validity. Third, evaluation procedures must adhere to the principle of parsimony; they must not only be feasible but they must be practicable within the parameters of the human resource system.

Each of these matters represent topics of considerable importance for personnel evaluation to be effectively planned and implemented. Adequate examination of all three is beyond the scope of a single paper of journal article length. The current paper will primarily focus on state legislative provisions and evaluation purposes. Brief attention will be given to the relation of technical soundness to evaluation purposes. More complete discussions of technical soundness and implementation issues (e.g., parsimony & practicability) are under development by the current authors as companion papers to the present document.

History of Educator Evaluation Statutory Development in Utah

Over the course of the past 19 years, various sessions of the Utah legislature have enacted four different and separate pieces of legislation that pertain to or include express or implied elements of educator evaluation. Each was enacted independently of the others and with seemingly different and distinctive purposes. Although all four acts have been subject to some degree of amendment since their initial enactment, they all continue in full force today.

The Utah Orderly Termination Procedures Act

The first of these statutes was passed in 1973 and is known as the Utah Orderly School Termination Procedures Act. Prior to that time, the State of Utah lacked any legislation governing educator termination except the following provision:

> Boards of education of local school districts may enter into written contracts for the employment of personnel for terms not to exceed five years, provided that nothing in the terms of such contracts shall restrict the power of such local boards to terminate such contracts for cause at any time (Utah Code Ann. §53-4-14 [1970]).

The original statutory purpose of the Orderly School Termination Procedures Act was: "to require school districts to adopt orderly termination procedures and to specify standards of due process and causes for termination" (Utah Code Ann. §53-51-2 [1981]). The due process guidelines specified in the statute include: a) notice of intent to terminate an employee; b) reason(s) for termination; c) opportunity for performance remediation; d) opportunity for a fair and impartial hearing including the right to counsel, to produce witnesses, to hear testimony, to cross-examine witnesses, and to examine documentary evidence. Although the Act outlines these due

process guidelines, each district is left the task of developing its own termination policy including specification of 'cause'.

The list of causes justifying termination vary from district to district, but virtually all identify incompetence or unprofessional conduct in one form or another. Because defensible dismissal charges generally require evidence based on some type of evaluation system, most (if not all) school districts have established and continue to maintain a method of evaluating performance that is linked to their policy and procedures governing orderly termination. Standards and procedures vary from district to district.

This statute departs from much of the national case law precedent in that no significant distinction (except the date of notice) is made between the due process rights of classified employees in educational support positions, year-to-year contract teachers ('non-tenured'), or continuing contract teachers ('tenured'). Thus, Utah school districts may assume that the full panoply of due process rights applies to virtually any employee.

Educational Professional Practices Act

Also in 1973, the Utah Legislature enacted a statute known as the Educational Professional Practices Act. This act established a state-wide professional practices commission composed of educators, at least half of whom must be classroom teachers. This commission was given, among other things, the following powers and duties:

(1) shall make recommendations to the State Board of Education and professional organizations of educators:
   a) concerning standards of professional performance, competence, and ethical conduct for persons holding certificates issued by the board;
   b) for the improvement of the education profession;

(2) shall adopt rules to carry out the purposes of this chapter;
(3) shall establish procedures for receiving and acting upon charges and recommendations regarding immoral, unprofessional, or incompetent conduct, unfitness for duty, or other violations of standards of ethical conduct, performance, and professional competence;
(4) shall establish the manner in which hearings are conducted and reported, and recommendations are submitted to the State Board of Education for its action;
(5) may
  (a) warn or reprimand a certificate holder;
  (b) recommend that the State Board of Education revoke or suspend a certificate;
  (c) enter into a written agreement requiring a current or former educator who has been the subject of a commission action to demonstrate to the satisfaction of the commission that the individual is rehabilitated and will conform to the standards of professional performance, competence, and ethical conduct; or
  (d) take other appropriate action; and
(6) may administer oaths, issue subpoenas, and make investigations relating to any matter before the commission (Utah Code Ann. §53A-7-110 [1991]).

Much of the language in this act suggests that its primary emphasis is on charges of undesirable behavior (e.g., unethical or immoral conduct) which might result in loss of licensure. However, inclusion of the terms "professional performance" and "competence" potentially broadens the scope of the act into instructional performance domains--domains typically evaluated by employing school districts.

In spite of the power and latitude given the Professional Practices Commission to investigate the competence of professional educators, nothing in the law prescribes or dictates evaluation standards or procedures. Indeed the law leaves to the commission the power "to adopt rules to carry out the purposes of this chapter" and to establish professional standards for the

entire state of educators. The commission has never adopted a fixed evaluation system or procedure, preferring to mold their investigation and evaluation on a case by case basis, regardless of the cause. Although obviously given broad authority including the right to recommend termination of a professional license, the legislature has never seen fit to link the powers and responsibilities accorded the professional practices commission with evaluation standards or procedures established in other pieces of legislation.

Career Ladder Act

In 1984 the Utah Legislature enacted the Career Ladder Act which was essentially a job enlargement, extended contract, and performance bonus provision. It had as its express purpose the following:

> The Legislature recognizes the importance of rewarding educators who strive to improve the quality of education, of providing incentives for educators employed by the public schools to continue to pursue excellence in education, of rewarding educators who demonstrate the achievement of excellence, and of properly compensating educators who assume additional educational responsibilities (Utah Code Ann. §53A-9-101 [1989]).

The legislature demanded that advancement on the career ladder program be contingent upon effective teaching performance. Evidence of effective teaching must include formal evaluation and assessment of student progress. Other criteria may include formal preparation and successful teaching experience. According to the law, the system is to provide "for frequent, comprehensive evaluation of teachers with less than 3 years' teaching experience, and periodic evaluations of other teachers."

Also according to the law, each school district must develop a program to evaluate its teachers for placement and

advancement on the career ladder. The statute prescribed that the evaluation procedure meet the following requirements:

(a) be fair, consistent, and valid according to generally accepted principles of personnel administration;
(b) incorporate clearly stated job descriptions;
(c) be in writing;
(d) involve teachers in the development of the evaluation instrument; and
(e) prior to any evaluation inform the teacher in writing about time frames in the evaluation procedure, the evaluation process, the types of criteria to be used in the evaluation and the factors to be evaluated and the procedure for requesting a review of the evaluation (Utah Code Ann. §53A-9-104 [1989]).

In 1988 the legislature enacted an amendment to the Career Ladder Act further requiring that district evaluation systems for placement and advancement on the career ladder be consistent with the Educator Evaluation Act which was enacted into law in 1987.

Educator Evaluation Act

The last of the four acts is known as the Educator Evaluation Act. Passed by the 1987 legislature its purposes are noted in the following statement of "Legislative Findings" which is a part of the act itself:

> The Legislature recognizes that the quality of public education can be improved and enhanced by providing for systematic, fair, and competent evaluation of public educators and remediation of those whose performance is inadequate. The desired purposes of evaluation are to allow the educator and the school district to promote the professional growth of the teacher, to identify and encourage teacher behaviors which contribute to student progress, to identify teachers according to their

abilities, and to improve the education system (Utah Code Ann. §53A-10-101 [1989]).

Under the provisions of this act each school district is to develop an evaluation program in consultation with a joint committee comprised of an equal number of classroom teachers and administrators appointed by the board. Any evaluation program adopted by a local board must provide or conform with the following standards:

(1) unless otherwise provided in the adopted program, the principal, the principal's designee, or the educator's immediate supervisor shall perform the educator evaluation;
(2) personal notice to the educator of the evaluation process at least 15 days prior to the first evaluation and receipt of a copy of the evaluation instrument, if an instrument is to be used;
(3) a reasonable number of observation periods for any evaluation to insure adequate opportunity for evaluation;
(4) the use of several types of evaluation and evidence, such as self-evaluation, student evaluation, peer evaluation, or systematic observations;
(5) that the educator may make a written response to all or any part of the evaluation and that the response will be attached to the evaluation;
(6) the evaluation will be a reliable and valid evaluation consistent with generally accepted professional standards for personnel evaluation systems; and
(7) within 15 days after the completed evaluation process the evaluation in writing shall be discussed with the educator. Following any revisions made after the discussion, a copy of the evaluation shall be filed in the educator's personnel file together with any related reports or documents. A copy of the evaluation and attachments shall be given to the educator (Utah Code Ann. §53A-10-106 [1991]).

When originally enacted, subsection 6 read, "the evaluation will be objective, reliable, and valid according to principles of effective instructional practice." The 1988 amendment has resulted in considerable disagreement in interpretation between the Utah Education Association and the Utah Association of School Superintendents. In 1990 an unsuccessful attempt was made in the legislature to remove the language requiring evaluation systems to be consistent with "professional standards for personnel evaluations systems."

The act further distinguishes between a "career educator", a "provisional educator", and a "probationary educator". Although the Utah legislature does not formally recognize tenure for educators, these terms approximate respectively the more typically recognizable terms of "tenured educator", "non-tenured educator", and an educator who has been advised that his/her performance is inadequate. This distinction is primarily important in the provisions for remediation and development in the Act. For instance, a probationary teacher must be provided "reasonable assistance" by the school district to improve performance. Further, provisional educators must be assigned a consulting teacher to "assist the provisional educator to become informed about the teaching profession and school system, but may not serve as an evaluator of the provisional teacher". The consulting teacher is required, where possible, to be a career teacher. This consulting teacher may receive additional compensation if his/her services require additional time "in the evaluation process", much like the job enhancement provisions in the Career Ladder Act.

Preliminary Considerations

Summary review of the four Utah statutory provisions suggests that each may have a **primary** purpose or legislative intent--independent of the other acts. However, certain provisions in some of these acts suggest secondary purposes which may intersect or be incompatible. For instance, the Career Ladder Act includes requirements regarding the evaluation of beginning teachers and others (see subsections 5 and 6 of the Utah Code

Ann. §53A-9-103 [1989]) that is only peripherally relevant to the primary purpose of the act. These provisions are more similar to the primary intent and language of the Educator Evaluation Act. Similarly, the Professional Practices Act is designed to give the responsibility for recommending "standards of professional performance, competence, and ethical conduct" to the Professional Practices Commission, yet the Educator Evaluation Act makes no reference to the Commission in its elaboration of evaluation standards and provisions.

Likewise, there is a lack of consistency in the definition of terms between the various acts. In particular, such terms as employee, educator, contract classified school employee, employee entitled/not entitled to employment in succeeding years, continuing contract employee, career educator, provisional teacher, beginning teacher, teacher with less than 3 years experience, and probationary teacher are not used consistently or clearly across the various legislative acts.

Thus overlapping purposes and provisions may present ambiguity and implementation dilemmas for school districts who seek to adopt evaluation policies and procedures to comply with each legislative act--**and** to adopt policies and procedures consistent with purposes of personnel evaluation. Similarly, ambiguity and inconsistency in terminology may leave school districts with questions regarding which acts apply to which employees in what regards. The net effect may be that local district evaluation policies and procedures may become unduly complicated, lack parsimony and practicability within the parameters of the human resource system.

Purposes, Types, and Technical Soundness of Evaluation

One of the most fundamental requirements for evaluation to be effective is the specification of the purposes for which assessment is undertaken. This has long been recognized in the

measurement literature and frequently identified as a problem when evaluation is encountering problems (Cronbach, 1990). Clearly articulating the purpose or purposes of evaluation is essential for several reasons including: 1) different types of evaluation are best applied for different purposes and personnel actions, and 2) components of the technical soundness of evaluation, notably validity, cannot be adequately addressed without the evaluation purpose.

Some Basic Types of Evaluation

Measurement and evaluation have been a major component of behavioral science for many decades. However, the maturation of evaluation was significantly enhanced by the work of a number of psychologists concerned about conceptual soundness, purposes, and uses as related to assessment (e.g., Anastasi, 1950, 1988; Bloom, Hastings, & Madaus, 1971; Drew, 1973; Haney, 1984; Gronlund & Linn, 1990). Several distinctions regarding types and uses of evaluation emerged from these works and have relevance to the current discussion.

Formative evaluation is evaluation that does not focus on terminal performance but rather on those behaviors or performance that represent the next step in a sequence. That sequence might refer to developing progressive skills that advance one from a novice teacher to a seasoned teacher. Formative evaluation aims at assisting the individual to improve specific performance areas such as classroom management techniques, or organizing instructional units logically for student learning. Thus, instrumentation is typically descriptive and diagnostic in nature. Formative evaluation is continuous, only quasi-formal, and because its purpose is performance enhancement, may have several foci relevant to the person's work in the context of the organization. It may identify areas of strength and limitation in the person being evaluated, the evaluation process, and the organization's structure and function. Typically, the person being evaluated is highly involved in the process including goal-setting, planning, and the evaluation itself (Harris & Monk, 1992).

Summative evaluation is quite different than formative. It refers to the assessment of terminal behaviors or overall performance which might result in some personnel action such as retention, contract renewal, promotion, or termination. Summative evaluation is used to make personnel decisions. It is more judgmental than formative evaluation and more global, taking a broader or more comprehensive view of overall performance--encompassing multiple performance domains such as instructional effectiveness, responsible work behaviors, cooperation with colleagues, and contributions to the classroom, school, and community. Summative evaluation is formal, relatively infrequent, and has a singular focus--the person being evaluated. The individual being evaluated is typically much less involved in the summative evaluation process than with formative evaluation, and may simply be informed of results and decisions. With this type of evaluation there must be a great deal of attention given to specific legal and contractual provisions (Harris & Monk, 1992).

Two additional measurement concepts are important but more commonly found in the general evaluation literature than that pertaining to personnel evaluation--norm- and criterion-referenced assessment. Norm-referenced evaluation is evaluation in which a person's performance is compared with that of other employees or with a group average. This is distinguished from criterion-referenced evaluation that compares a person's performance to some specific established standard (the criterion); his or her performance is not compared with that of other people. Often criterion-referenced evaluation attempts to assess specific skill areas rather than generating some global or composite measure as is the case with norm-referenced assessment. Each of these types of evaluation might best be applied for different personnel evaluation purposes.

Matching Evaluation Types to Purposes

There are a multitude of different reasons for undertaking personnel evaluation. Let's explore for a moment a selected few and the types of evaluation that might be most appropriately used.

Staff development is a personnel function which exemplifies a desire to upgrade or improve skills among a group of educators. This purpose might be best evaluated using formative evaluation that is criterion referenced, particularly if one is pinpointing specific skill areas. Norm-referenced assessment might be helpful if the perspective is a competitive one, but the predominant and most useful approach would be criterion-referenced formative assessment since this combination addresses on-going evaluation that focuses on specific skills.

Rewarding performance is another personnel-related action that requires evaluation--as in the case of merit pay decisions. Evaluation used for rewarding performance might include at least three of the four evaluation types noted. Norm-referenced evaluation would likely be used when rewarding someone for being more able than others (thereby comparing his or her performance with that of others, or some norm or average). By contrast, criterion-referenced evaluation might be employed when rewarding all persons who reach a certain level of skill. Lastly, if one is intending to reward improvement then it is necessary to track the progressive development of performance using formative evaluation (assuming that it is undesirable to wait an extended period of time to reinforce good performance).

Promotion is a personnel action that is typically related to the cumulative performance record or accomplishments (here we are not addressing promotion that occurs because of longevity). Evaluation used for promotion might employ the approaches noted for rewarding performance. Additionally there may be a summative assessment component in cases in which promotion reflects a significant movement from one level to another.

Retention, contract renewal, or termination are personnel actions that require a different approach. These personnel decisions typically rest on a criterion-referenced assessment to determine if an individual did or did not demonstrate the minimum job skills necessary to continue employment. The final decision is based on formal, summative evaluation and, as noted earlier, may have contractual and legal parameters such as due process.

It is possible, however, that norm-referenced evaluation would be used in some termination decisions, such as reduction-in-force actions. The type of evaluation used for these personnel actions may also vary with the stage of the decision process. For instance, prior to a final employment decision there should be efforts to remediate sub-standard performance. This evaluation may become focused on skill areas needing improvement and therefore employ formative evaluation.

Evaluation Purposes and Technical Soundness

One area of personnel evaluation that has continually been a concern relates to accuracy. The Joint Committee on Standards for Educational Evaluation (1988) identified accuracy as one of its major categories of standards aimed at improving personnel evaluation. Accuracy is central to the technical soundness of measurement and verified by the validity and reliability of assessment. Although there are many technical components of validity and reliability, one of the fundamental issues often overlooked is the relation of these measurement constructs to evaluation purposes.

Validity is a term that is often used broadly or superficially. However, validity is a term that should not be used without a modifier. With regard to measurement validity, one must always ask, "valid for what"? Validity cannot be adequately established without a contextual referent. For instance, an evaluation instrument or procedure that is valid for screening applicants for initial employment suggests content validity and predictive validity are appropriate. That is, the screening instrument's items should be clearly job-relevant (content validity); and, there should be a clear relationship between an applicant's performance on the screening instrument and his/her subsequent performance on the job (predictive validity). Whereas, rewarding teachers that have made exemplary contributions to student learning may suggest that criterion-related validity and discriminant validity are appropriate. That is, there should be a clear relationship between a teacher's performance and some indicator of student learning such as achievement test scores (criterion-related validity); and the measure of teacher performance should adequately "discriminate" between "exemplary" performance and adequate performance (discriminant validity). Thus, an evaluation instrument is not just valid, but valid for a specific purpose; validity is not established in a vacuum.

Reliability of evaluation reflects assessment accuracy, consistency, or stability (freedom from error). Reliability refers to

how accurately a given measurement provides evidence of change when there is change and does not indicate change when there is none. There is a close relationship between reliability and validity; it is necessary to know what one is evaluating to determine how accurately it is being measured. In the teacher evaluation illustration noted above, it is necessary to clearly define the criterion term. Without a clear definition of "exemplary contributions to student learning", one cannot verify the stability (reliability) of measurement.

Many administrators and legislators focus immediate concern on validity and reliability when examining personnel evaluation. These are legitimate topics and should be considered carefully. They are not, however, the first matters that need to be addressed. Fundamental to validity and reliability are issues of evaluation purpose. In fact, a clear articulation of the purpose of evaluation is the very foundation upon which the effectiveness of any assessment system rests.

Congruency Between Legislative Intent and Evaluation Purposes

Most of the legislation reviewed in this paper only addresses evaluation purpose sparingly, at least in terms of specific language. As one would expect, the **Educator Evaluation Act** provides the most attention to purposes, with the following being noted: 1) to promote teacher growth, 2) to identify and encourage teacher behaviors that contribute to student progress, 3) to identify teachers according to their abilities, and 4) to improve the education system.

Viewed broadly the Educator Evaluation Act appears to have a staff development purpose. From the discussion above we know that some formative, criterion-referenced assessment is probably most appropriate. However, there are several points in the language of the law that run counter to typical formative evaluation procedures. First of all, the process is not continuous

although local boards are to provide for such evaluation at least twice each school year (Utah Code Ann. §53A-10-104 [1989]). Secondly, the people involved in the evaluation are not necessarily those we might expect in effective formative evaluation. The specified evaluation process is heavily management-oriented with the principal orienting all educators regarding the evaluation program (Utah Code Ann. §53A-10-105 [1989]). Additionally, the principal, his or her designee, or the immediate supervisor is suggested as the primary evaluator (Utah Code Ann. §53A-10-106 [1991]). No specific mention is made of involving the person being evaluated in the process, a factor generally accepted as appropriate in the application of formative evaluation to personnel matters (Harris & Monk, 1992). Further, consulting teachers assigned to provisional teachers are not allowed to serve as evaluators (Utah Code Ann. §53A-10-108 [1991]). Finally, language in later portions of the law sounds very much like a formal process leading to termination (e.g., Utah Code Ann. §53A-10-109 [1989] on final evaluation and Utah Code Ann. §53A-10-110 [1991] on review of evaluation -- time limit on request). Thus although the Educator Evaluation Act seems to intend staff development, the purpose appears to become confused because the methods specified are not consistent with evaluation that is appropriate for staff development.

The **Career Ladder Act** has stated purposes of rewarding performance excellence as well as providing additional compensation for those who assume additional responsibilities. Although these are two distinct purposes and should involve different evaluation methods, they are often conceptually interrelated in some school district career ladder policies (Hart, 1987, 1990). To some degree, one has become the measure of the other--that is, the assumption of additional responsibilities has become the criterion of excellence. This appears to confuse or certainly make indistinct the two purposes and their corresponding evaluation methods. To further complicate matters, assigning (or allowing) additional responsibilities has become a means for rewarding excellence in some cases. One can only surmise that it is easier to count extra days or hours worked and compute per

diem pay than it is to evaluate and reward excellent quality of performance as a teacher.

Implementation practices aside, the two stated purposes suggest certain evaluation approaches. If we are to reward excellent performance, we must have a means of defining excellence. Because excellence was not defined in the law, school districts have significant latitude to determine an operational definition. Excellence may have elements of both norm- and criterion-referencing. To be excellent suggests attainment of a high quality standard, thus implying a criterion-referenced assessment. However, excellent can also mean to be of superior merit--that is, more able than others, thus implying a norm-referenced assessment. The law does not provide much guidance regarding formative vs. summative evaluation. However, if the reward for excellence is additional compensation, one could argue that a property interest may be at stake, thus requiring a summative evaluation to make that personnel decision.

The second purpose of the Career Ladder Act relates to additional compensation for additional responsibility. Interpreted simply, this is a relatively straightforward matter of extra pay for extra work and requires little consideration for performance evaluation. As suggested earlier, in practice this purpose has become entangled with that of rewarding excellence, perhaps because of difficulties or burdens in evaluating performance.

As described earlier, the **Educational Professional Practices Act** established a state-wide professional practices commission empowered to undertake a number of matters. Among these responsibilities are certain actions that relate to personnel evaluation including recommendations concerning standards of professional performance, competence, and ethical conduct. Although the law does not specifically speak to evaluation, it is difficult to establish meaningful standards of professional performance and competence in the absence of some attention to how those performance and competence standards can be assessed.

For the most part this act seems to address causes that lie in the area of undesirable behavior (e.g. unethical, immoral, or illegal conduct) rather than instructional performance. However, the term incompetence is also included, potentially broadening the scope of the act into areas related to instructional performance. The act's reference to incompetence implies attention to evaluation of instructional performance. Efforts to improve such performance or to provide remediation would most often employ formative, criterion-referenced evaluation; or, in the case of retention or termination decisions, summative evaluation methods would be employed. However, the Professional Practices Commission has largely chosen not to deal with questions of instructional performance. Due to limited time and resources, complaints from parent and other client groups regarding instructional issues are commonly forwarded to the local school districts for action. The commission does track the outcomes of these referrals.

Action by the commission is more typically aimed at alleged ethical and/or behavioral misconduct, especially misconduct which might result in loss of licensure. The majority of its action is precipitated by required law enforcement reports (Utah Code Ann. §77-26-22 [1989]) regarding alleged sexual misconduct or other behavioral offenses regulated under §58-37-8 or in Title 76 of the Utah Code. The focus on aberrant or undesirable behavior raises questions regarding the appropriateness of norm-referenced versus criterion-referenced evaluation. With the exception of some very basic minimums in terms of routine demeanor, the desirability of behavior is nearly always determined by reference to norms. Those norms may vary to some degree because local standards often play a prominent role. However, more broadly based professional norms or legal provisions may also be used to establish a standard of conduct. Although ethical and legal standards may suggest criterion-referenced assessment, determination of those standards are clearly norm-referenced.

The **Orderly School Termination Procedures Act** clearly involves termination as its purpose although it also includes an element pertaining to remediation. The Orderly Termination Act

suggests that the main purpose is a requirement of due process and specification of cause for termination. All evaluation leading to termination will need to be summative for the reasons outlined earlier. The summative evaluation and final termination decision will most appropriately be preceded by a remediation element employing formative evaluation procedures.

Summary Comments and Recommendations

Analysis of Utah's four major pieces of legislation related to educator evaluation illustrates a number of difficulties that are also evident in other parts of the country (Claxton, 1986). These difficulties may be due, in part, to the evolutionary nature of educator evaluation as well as statutory compromise that results from the political and legislative process.

One major difficulty may be inherent in legislative language. The legislative purpose of an act may be only vaguely stated, if at all, and may not clearly or succinctly translate into purposes for personnel evaluation. Further, legislative language, particularly definitions of terms, may vary between and among related acts. Thus, how performance or competence is to be assessed--as well as whose performance is to be assessed by whom--becomes unclear, subject to multiple interpretations, and a possible source of conflict in local school districts.

A second problem is the lack of independence of purpose between and among various evaluation-related statutes. Although the primary purpose of each act may be clearly distinct from that of the other acts, the secondary purposes may intersect or be incompatible with one another. This "purpose contamination" may reduce the simplicity and ease of evaluation implementation for local school districts, as well as reduce the effectiveness of evaluation for specific purposes.

Similarly, mixing different purposes **within** the same legislative provision--purposes that may not be compatible with respect to effective personnel evaluation methods--may hamper

implementation efforts. This complication may compromise the effectiveness of evaluation implementation efforts in local school districts by creating complicated and unmanageable systems of evaluation that lack parsimony and practicability.

Legislation also seems to combine attention to educational or instructional performance matters with behavioral misconduct matters. The term competence is frequently used in a fashion that can apply to either of these areas. However, in practice, these employee behaviors often reflect different domains of personnel performance; and, unless one is influencing the other (e.g., aberrant behavior affecting instructional performance), they require different purposes and methods of evaluation.

In sum, because the purpose of evaluation is foundational to the methods, technical soundness, and effectiveness of an evaluation system, it is critical that these purposes be congruent with the legislative intent of educator evaluation statutes. The Utah case study illustrates some of the evaluation complications that can occur when the legislation lacks clarity, simplicity, and compatibility with regard to purposes of evaluation.

Recommendations

The following recommendations are made to help administrators and policy-makers address some of the complications and ambiguities related to educator evaluation in Utah. The first set of recommendations is addressed to local school administrators who must implement appropriate educator evaluation practices under existing Utah laws. The second set of recommendations is addressed to state policy-makers who may wish to resolve some of the current educator evaluation complications by revising existing legislative acts.

To Administrators:

1. Keep purposes of each and every evaluation clear. (i.e., is it formative, summative, or both; is it to promote professional growth, to determine eligibility for career ladder monies, to

be used for possible termination action, or some combination?)

2. Match an evaluation or an evaluation system to the intended purpose(s).

3. Make a "good faith effort" to identify and comply with all procedural and substantive requirements noted in the Utah Orderly Termination Procedures Act, the Career Ladder Act and the Educator Evaluation Act.

4. Because current legislative language requires "a reliable and valid evaluation consistent with generally accepted professional standards for personnel evaluation systems" and given the fact that no single set of standards enjoys widespread acceptance, local school boards should adopt a set of "defensible" standards which enjoys the support of both local teachers and administrators.

5. Establish and maintain a continuous training and monitoring system of all evaluation instruments and procedures. Data need to be gathered and evaluated constantly to preserve the integrity of the system. This will require the use of highly trained specialized personnel.

To Policy-Makers:

1. Incorporate critical elements and primary purposes of the four evaluation-related acts into an omnibus educator evaluation act to reduce the ambiguity, overlap, and potential contradictions currently created by terminology and provisions of the current four separate acts. For instance, the omnibus act could include:

   a) a provision for formative evaluation and staff development (similar to some aspects of the existing Educator Evaluation Act);
   b) a provision for summative evaluation or employment retention/termination decisions including due process

provisions (similar to some aspects of the Educator Evaluation Act and primary aspects of the Orderly Termination Act);

c) a provision for rewarding performance (similar to primary aspects of the Career Ladder Act);

d) a provision for revocation of licensure (a primary aspect of the Educational Professional Practices Act).

Most importantly, this omnibus act should have clear and consistent use of terms and make provision for the full array of personnel functions and decisions that represent the various purposes for educator evaluation.

2. The degree of specificity in a newly created omnibus act should be somewhat less than currently dictated by similar acts. In particular, the specific methods, time-frames, etc. that are appropriate for each major type of evaluation should be determined by a joint committee of teachers and administrators at the local district level. The omnibus legislation should provide the basic purposes and broad legal principles that must be implemented. The local districts should be granted the discretion to adopt implementation policies and practices that fit local district needs and yet adhere to broad legal principles of fairness, accuracy, effectiveness, and practicability.

References

Anastasi, A. (1950). The concept of validity in the interpretation of test scores. Educational and Psychological Measurement, 10, 67-78.

Anastasi, A. (1988). Psychological testing (6th ed.). New York: Macmillan.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill.

Claxton, W. P. (1986). Remediation: The evolving fairness in teacher dismissal. Journal of Law and Education, 15, 181-193.

Cronbach, L.J. (1990). Essentials of psychological testing (5th ed.). New York: Harper & Row.

Drew, C. J. (1973). The use of criterion-referenced and normative-referenced measures in assessing minority group children. Journal of School Psychology, 11, 323-329.

Gronlund, N. E., & Linn, R. L. (1990). Measurement and evaluation in teaching. New York: Macmillan.

Haney, W. (1984). testing and reasoning and reasoning about testing. Review of Educational Research, 54, 597-654.

Harris, B. M., & Monk, B. J. (1992). Personnel administration in education: Leadership for instructional improvement (3rd ed.), Boston: Allyn & Bacon.

Hart, A. W. (1987). A career ladder's effect on teacher career and work attitudes. American Educational Research Journal, 24, 4, 479-503.

Hart, A. W. (1990). Impacts of the school social unit on teacher authority during work redesign. American Educational Research Journal, 27 (3), 503-532.

Joint Committee on Standards for Educational Evaluation. (1988). The Personnel Evaluation Standards: How to Assess Systems for Evaluating Educators. Newbury Park, CA: Sage.

Salvia, J., & Ysseldyke, J. E. (1985). Assessment in special and remedial education (3rd ed.). Boston: Houghton Mifflin.

Utah Code Annotated §53-4-14 (Supp.1970).

Utah Code Annotated §53-51-2 (Supp.1981). Utah Orderly Termination Procedures Act.

Utah Code Annotated §§53A-9-101,103 (1989). Career Ladder Act.

Utah Code Annotated §§53A-10-101,104,105,109 (1989). Educator Evaluation Act.

Utah Code Annotated §77-26-22 (1989). Criminal Identification.

Utah Code Annotated §§53A-10-106,108,110 (Supp.1991). Educator Evaluation Act.

Utah Code Annotated. §53A-7-110 (Supp.1991). Educational Professional Practices Act.